

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Huse, William D.
- (ii) TITLE OF INVENTION: SURFACE EXPRESSION LIBRARIES OF  
RANDOMIZED PEPTIDES
- (iii) NUMBER OF SEQUENCES: 61
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
  - (B) STREET: 444 South Flower Street, Suite 2000
  - (C) CITY: Los Angeles
  - (D) STATE: California
  - (E) COUNTRY: United States
  - (F) ZIP: 90071
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Campbell, Cathryn A
  - (B) REGISTRATION NUMBER: 31,815
  - (C) REFERENCE/DOCKET NUMBER: P31 9072
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (619) 535-9001
  - (B) TELEFAX: (619) 535-8949

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7294 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: both
  - (D) TOPOLOGY: circular

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA	180

GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA	240
TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCCGGTCT GGTTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG	360
TCTTTCGGGC TTCCTCTTAA TCTTTTGTAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGCAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT	600
GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT	660
AATTCCTTTT GGC GTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT	780
TCTTCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA	840
CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTT	900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG	960
AATATCCGGT TCTTGTCAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC	1020
TGTACACCGT TCATCTGTCC TCTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC	1080
GTCTGCGCCT CGTTCCGGCT AAGTAACATG GAGCAGGTCG CGGATTTCGA CACAATTTAT	1140
CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT	1200
CAAAGATGAG TGTTTTAGTG TATTCTTTCG CCTCTTTCGT TTTAGGTTGG TGCCTTCGTA	1260
GTGGCATTAC GTATTTTACC CGTTTAATGG AAACCTCCTC ATGAAAAAGT CTTTAGTCCT	1320
CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA	1380
CGATCCCGCA AAAGCGGCCT TTAACCTCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA	1440
TGCGTGGGCG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTTAAGAA	1500
ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT	1560
TTTTTGAGAG TTTTCAACGT GAAAAAATTA TTATTCGCAA TTCCTTTAGT TGTTCTTTT	1620
TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA	1680
TTTACTAACG TCTGGAAAGA CGACAAAAC TTAGATCGTT ACGCTAACTA TGAGGGTTGT	1740
CTGTGGAATG CTACAGGCGT TGAGTTTGT ACTGGTGACG AAACCTCAGT TTACGGTACA	1800
TGGGTTCCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT	1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT	1920
ATTCGGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA	1980

AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCATGTTT	2040
CAGAATAATA GGTTCGAAA TAGGCAGGG GCATTAAC TG TTTATACGGG CACTGTTACT	2100
CAAGGCACTG ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG	2160
TATGACGCTT ACTGGAACGG TAAATTCAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA	2220
GATCCATTTCG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT	2280
GCTGGCGGCG GCTCTGGTGG TGGTCTGGT GGC GGCTCTG AGGGTGGTGG CTCTGAGGGT	2340
GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GGCGGTTCCG GTGGTGGCTC TGGTTCGGT	2400
GATTTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT	2460
GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAAC TTGATT CTGTCGCTAC TGATTACGGT	2520
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT	2580
GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT	2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT	2700
TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA	2760
TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG	2820
TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCGGT	2880
TATTATTGCG TTTCCTCGGT TTCCTTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC	2940
TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CTATTTTCATT GTTTCTTGCT CTTATTATTG	3000
GGCTTAACTC AATTCTTG TG GGTATCTCT CTGATATTAG CGCTCAATTA CCCTCTGACT	3060
TTGTTTCAGGG TGTTTCAGTTA ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TATGTTATTC	3120
TCTCTGTAAA GGCTGCTATT TTCATTTTTG ACGTTAAACA AAAAATCGTT TCTTATTTGG	3180
ATTGGGATAA ATAATATGGC TGTTTATTTT GTAAC TGGCA AATTAGGCTC TGGAAAGACG	3240
CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT	3300
CTTGATTTAA GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAAC GCCTCGCGTT	3360
CTTAGAATAC CGGATAAGCC TTCTATATCT GATTTGCTTG CTATTGGGCG CGGTAATGAT	3420
TCCTACGATG AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTGCGGTAC TTGGTTTAAT	3480
ACCCGTTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTTCT ACATGCTCGT	3540
AAATTAGGAT GGGATATTAT CTCCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG	3600
CGTTCTGCAT TAGCTGAACA TGTTGTTTAT TGTCGTCGTC TGGACAGAAT TACTTTACCT	3660
TTTGTCGGTA CTTTATATTC TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT	3720
GTTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT	3780

ACTGGTAAGA ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTTCTAG TAATTATGAT	3840
TCCGGTGTTT ATTCTTATTT AACGCCTTAT TTATCACACG GTCGGTATTT CAAACCATTA	3900
AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTTGA AAAAGTTTTT CACGCTTCTT	3960
TGTCTTGCGA TTGGATTGTC ATCAGCATT ACATATAGTT ATATAACCCA ACCTAAGCCG	4020
GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTTGATA AATTCATAT TGA CTCTCT	4080
CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT	4140
AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTTATG TACTGTTTCC	4200
ATTAAAAAGG TAATTCAAAT GAAATTGTTA AATGTAATTA ATTTTGTTTT CTGTATGTTT	4260
GTTTCATCAT CTCTTTTGC TCAGGTAATT GAAATGAATA ATTCGCCTCT GCGCGATTTT	4320
GTAACCTGGT ATTCAAAGCA ATCAGGCGAA TCCGTTATTG TTTCTCCCGA TGTAAGGTT	4380
ACTGTTACTG TATATTCATC TGACGTTAAA CCTGAAAATC TACGCAATTT CTTTATTTCT	4440
GTTTTACGTG CTAATAATTT TGATATGGTT GGTTCATTC CTTCATTAT TTAGAAGTAT	4500
AATCCAAACA ATCAGGATTA TATTGATGAA TTGCCATCAT CTGATAATCA GGAATATGAT	4560
GATAATTCCG CTCCTTCTGG TGGTTTCTTT GTTCCGCAA ATGATAATGT TACTCAAAC	4620
TTTAAATTA ATAACGTTTC GGCAAAGGAT TTAATACGAG TTGTCGAATT GTTTGTAAAG	4680
TCTAATACTT CTAAATCCTC AAATGTATTA TCTATTGACG GCTCTAATCT ATTAGTTGTT	4740
AGTGCACCTA AAGATATTTT AGATAACCTT CCTCAATTCC TTTCTACTGT TGATTTGCCA	4800
ACTGACCAGA TATTGATTGA GGGTTTGATA TTTGAGGTT AGCAAGGTGA TGCTTTAGAT	4860
TTTTCATTTG CTGCTGGCTC TCAGCGTGGC ACTGTTGCAG GCGGTGTTAA TACTGACCGC	4920
CTCACCTCTG TTTTATCTTC TGCTGGTGGT TCGTTCGGTA TTTTAAATGG CGATGTTTTA	4980
GGGCTATCAG TTCGCGCATT AAAGACTAAT AGCCATTCAA AAATATTGTC TGTGCCACGT	5040
ATTCTTACGC TTTCAGGTCA GAAGGGTTCT ATCTCTGTTG GCCAGAATGT CCCTTTTATT	5100
ACTGGTCGTG TGA CTGGTGA ATCTGCCAAT GTAAATAATC CATTCAGAC GATTGAGCGT	5160
CAAAATGTAG GTATTTCCAT GAGCGTTTTT CTGTTGCAA TGGCTGGCGG TAATATTGTT	5220
CTGGATATTA CCAGCAAGGC CGATAGTTTG AGTTCTTCTA CTCAGGCAAG TGATGTTATT	5280
ACTAATCAAA GAAGTATTGC TACAACGGTT AATTTGCGTG ATGGACAGAC TCTTTTACTC	5340
GGTGGCCTCA CTGATTATAA AAACACTTCT CAAGATTCTG GCGTACCGTT CCTGTCTAAA	5400
ATCCCTTTAA TCGGCCTCCT GTTTAGCTCC CGCTCTGATT CCAACGAGGA AAGCACGTTA	5460
TACGTGCTCG TCAAAGCAAC CATAGTACGC GCCCTGTAGC GGCGCATTA GCGCGGCGGG	5520
TGTGGTGGTT ACGCGCAGCG TGACCGCTAC ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT	5580

CGCTTTCTTC	CCTTCCTTTC	TCGCCACGTT	CGCCGGCTTT	CCCCGTCAAG	CTCTAAATCG	5640
GGGGCTCCCT	TTAGGGTTCC	GATTTAGTGC	TTTACGGCAC	CTCGACCCCA	AAAAACTTGA	5700
TTTGGGTGAT	GGTTCACGTA	GTGGGGCATC	GCCCTGATAG	ACGGTTTTTC	GCCCTTTGAC	5760
GTTGGAGTCC	ACGTTCTTTA	ATAGTGGACT	CTTGTTCCAA	ACTGGAACAA	CACTCAACCC	5820
TATCTCGGGC	TATTCTTTTG	ATTTATAAGG	GATTTTGCCG	ATTTGGAAC	CACCATCAAA	5880
CAGGATTTTC	GCCTGCTGGG	GCAAACCAGC	GTGGACCGCT	TGCTGCAACT	CTCTCAGGGC	5940
CAGGCGGTGA	AGGGCAATCA	GCTGTTGCCC	GTCTCGCTGG	TGAAAAGAAA	AACCACCCTG	6000
GCGCCCAATA	CGCAAACCGC	CTCTCCCCGC	GCGTTGGCCG	ATTCATTAAT	GCAGCTGGCA	6060
CGACAGGTTT	CCCGACTGGA	AAGCGGGCAG	TGAGCGCAAC	GCAATTAATG	TGAGTTAGCT	6120
CACTCATTAG	GCACCCCAGG	CTTTACACTT	TATGCTTCCG	GCTCGTATGT	TGTGTGGAAT	6180
TGTGAGCGGA	TAACAATTTT	ACACAGGAAA	CAGCTATGAC	CAGGATGTAC	GAATTCGCAG	6240
GTAGGAGAGC	TCGGCGGATC	CTAGGCTGAA	GGCGATGACC	CTGCTAAGGC	TGCATTCAAT	6300
AGTTTACAGG	CAAGTGCTAC	TGAGTACATT	GGCTACGCTT	GGGCTATGGT	AGTAGTTATA	6360
GTTGGTGCTA	CCATAGGGAT	TAAATTATTC	AAAAAGTTTA	CGAGCAAGGC	TTCTTAACCA	6420
GCTGGCGTAA	TAGCGAAGAG	GCCCCGACCG	ATCGCCCTTC	CCAACAGTTG	CGCAGCCTGA	6480
ATGGCGAATG	GCGCTTTGCC	TGGTTTCCGG	CACCAGAAGC	GGTGCCGGAA	AGCTGGCTGG	6540
AGTGCGATCT	TCCTGAGGCC	GATACGGTCG	TCGTCCCCTC	AAACTGGCAG	ATGCACGGTT	6600
ACGATGCGCC	CATCTACACC	AACGTAACCT	ATCCCATTAC	GGTCAATCCG	CCGTTTGTTT	6660
CCACGGAGAA	TCCGACGGGT	TGTTACTCGC	TCACATTTAA	TGTTGATGAA	AGCTGGCTAC	6720
AGGAAGGCCA	GACGCGAATT	ATTTTGTATG	GCGTTCCTAT	TGGTTAAAAA	ATGAGCTGAT	6780
TTAACAAAAA	TTTAACGCGA	ATTTTAACAA	AATATTAACG	TTTACAATTT	AAATATTTGC	6840
TTATACAATC	TTCTGTTTTT	TGGGGCTTTT	CTGATTATCA	ACCGGGGTAC	ATATGATTGA	6900
CATGCTAGTT	TTACGATTAC	CGTTCATCGA	TTCTCTTGTT	TGCTCCAGAC	TCTCAGGCAA	6960
TGACCTGATA	GCCTTTGTAG	ATCTCTCAAA	AATAGCTACC	CTCTCCGGCA	TTAATTTATC	7020
AGCTAGAACG	GTTGAATATC	ATATTGATGG	TGATTTGACT	GTCTCCGGCC	TTTCTCACCC	7080
TTTTGAATCT	TTACCTACAC	ATTACTCAGG	CATTGCATTT	AAAATATATG	AGGGTTCTAA	7140
AAATTTTTTAT	CCTTGCGTTG	AAATAAAGGC	TTCTCCCGCA	AAAGTATTAC	AGGGTCATAA	7200
TGTTTTTGGT	ACAACCGATT	TAGCTTTATG	CTCTGAGGCT	TTATTGCTTA	ATTTTGCTAA	7260
TTCTTTGCCT	TGCCTGTATG	ATTTATTGGA	CGTT			7294

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7320 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCGCAGA ATTGGGAATC AACTGTTACA TGAATGAAA CTTCCAGACA CCGTACTTTA	180
GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA	240
TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCCGGTCT GGTTTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG	360
TCTTTCGGGC TTCCTCTTAA TCTTTTGTAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGCAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT	600
GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT	660
AATTCCTTTT GGC GTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT	780
TCTTCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA	840
CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTT	900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG	960
AATATCCGGT TCTTGCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC	1020
TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC	1080
GTCTGCGCCT CGTTCCGGCT AAGTAACATG GAGCAGGTCG CGGATTTCTGA CACAATTTAT	1140
CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT	1200
CAAAGATGAG TGTTTTAGTG TATTCTTTTCG CCTCTTTTCGT TTTAGGTTGG TGCCTTCGTA	1260
GTGGCATTAC GTATTTTACC CGTTTAATGG AAACCTCCTC ATGAAAAAGT CTTTAGTCCT	1320
CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA	1380
CGATCCCGCA AAAGCGGCCT TTAACCTCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA	1440

TGCGTGGGCG	ATGGTTGTTG	TCATTGTCCG	CGCAACTATC	GGTATCAAGC	TGTTTAAGAA	1500
ATTCACCTCG	AAAGCAAGCT	GATAAACCGA	TACAATTAAA	GGCTCCTTTT	GGAGCCTTTT	1560
TTTTTGGAGA	TTTCAACGT	GAAAAAATTA	TTATTCGCAA	TTCCTTTAGT	TGTTCCCTTC	1620
TATTCTCACT	CCGCTGAAAC	TGTTGAAAGT	TGTTTAGCAA	AACCCCATAC	AGAAAATTCA	1680
TTTACTAACG	TCTGGAAAGA	CGACAAAAC	TTAGATCGTT	ACGCTAACTA	TGAGGGTTGT	1740
CTGTGGAATG	CTACAGGCGT	TGTAGTTTGT	ACTGGTGACG	AAACTCAGTG	TTACGGTACA	1800
TGGGTTCCCTA	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGCGGT	1860
TCTGAGGGTG	GCGGTTCTGA	GGGTGGCGGT	ACTAAACCTC	CTGAGTACGG	TGATACACCT	1920
ATTCCGGGCT	ATACTTATAT	CAACCCTCTC	GACGGCACTT	ATCCGCCTGG	TACTGAGCAA	1980
AACCCCGCTA	ATCCTAATCC	TTCTCTTGAG	GAGTCTCAGC	CTCTTAATAC	TTTCATGTTT	2040
CAGAATAATA	GGTTCGAAA	TAGGCAGGGG	GCATTAAC	TTTATACGGG	CACTGTTACT	2100
CAAGGCACTG	ACCCCGTTAA	AACTTATTAC	CAGTACACTC	CTGTATCATC	AAAAGCCATG	2160
TATGACGCTT	ACTGGAACGG	TAAATTCAGA	GACTGCGCTT	TCCATTCTGG	CTTTAATGAA	2220
GATCCATTCG	TTTGTGAATA	TCAAGGCCAA	TCGTCTGACC	TGCCTCAACC	TCCTGTCAAT	2280
GCTGGCGGCG	GCTCTGGTGG	TGGTTCTGGT	GGCGGCTCTG	AGGGTGGTGG	CTCTGAGGGT	2340
GGCGGTTCTG	AGGGTGGCGG	CTCTGAGGGA	GGCGGTTCCG	GTGGTGGCTC	TGGTTCCGGT	2400
GATTTTGATT	ATGAAAAGAT	GGCAAACGCT	AATAAGGGGG	CTATGACCGA	AAATGCCGAT	2460
GAAAACGCGC	TACAGTCTGA	CGCTAAAGGC	AAACTTGATT	CTGTCGCTAC	TGATTACGGT	2520
GCTGCTATCG	ATGGTTTCAT	TGGTGACGTT	TCCGGCCTTG	CTAATGGTAA	TGGTGCTACT	2580
GGTGATTTTG	CTGGCTCTAA	TTCCCAAATG	GCTCAAGTCG	GTGACGGTGA	TAATTCACCT	2640
TTAATGAATA	ATTTCCGTCA	ATATTTACCT	TCCCTCCCTC	AATCGGTTGA	ATGTCGCCCT	2700
TTTGTCTTTA	GCGCTGGTAA	ACCATATGAA	TTTTCTATTG	ATTGTGACAA	AATAAACTTA	2760
TTCCGTGGTG	TCTTTGCGTT	TCTTTTATAT	GTTGCCACCT	TTATGTATGT	ATTTTCTACG	2820
TTTGCTAACA	TACTGCGTAA	TAAGGAGTCT	TAATCATGCC	AGTTCTTTTG	GGTATTCGGT	2880
TATTATTGCG	TTTCCTCGGT	TTCCCTTCTG	TAACTTTGTT	CGGCTATCTG	CTTACTTTTC	2940
TTAAAAAGGG	CTTCGGTAAG	ATAGCTATTG	CTATTTTCATT	GTTTCTTGCT	CTTATTATTG	3000
GGCTTAACTC	AATTCTTG	GGTTATCTCT	CTGATATTAG	CGCTCAATTA	CCCTCTGACT	3060
TTGTTTCAGG	TGTTTCAGTTA	ATTCTCCCGT	CTAATGCGCT	TCCCTGTTTT	TATGTTATTC	3120
TCTCTGTAAA	GGCTGCTATT	TTCATTTTTG	ACGTTAAACA	AAAAATCGTT	TCTTATTTGG	3180
ATTGGGATAA	ATAATATGGC	TGTTTATTTT	GTAAGTGGCA	AATTAGGCTC	TGAAAGACG	3240

CTCGTTAGCG	TTGGTAAGAT	TTAGGATAAA	ATTGTAGCTG	GGTGCAAAAT	AGCAACTAAT	3300
CTTGATTTAA	GGCTTCAAAA	CCTCCCGCAA	GTCGGGAGGT	TCGCTAAAC	GCCTCGCGTT	3360
CTTAGAATAC	CGGATAAGCC	TTCTATATCT	GATTTGCTTG	CTATTGGGCG	CGGTAATGAT	3420
TCCTACGATG	AAAATAAAAA	CGGCTTGCTT	GTTCTCGATG	AGTGCGGTAC	TTGGTTTAAT	3480
ACCCGTTCTT	GGAATGATAA	GGAAAGACAG	CCGATTATTG	ATTGGTTTCT	ACATGCTCGT	3540
AAATTAGGAT	GGGATATTAT	CTTCCTTGTT	CAGGACTTAT	CTATTGTTGA	TAAACAGGCG	3600
CGTTCTGCAT	TAGCTGAACA	TGTTGTTTAT	TGTCGTCGTC	TGGACAGAAT	TACTTTACCT	3660
TTTGTCGGTA	CTTTATATTC	TCTTATTACT	GGCTCGAAAA	TGCCTCTGCC	TAAATTACAT	3720
GTTGGCGTTG	TTAAATATGG	CGATTCTCAA	TTAAGCCCTA	CTGTTGAGCG	TTGGCTTTAT	3780
ACTGGTAAGA	ATTTGTATAA	CGCATATGAT	ACTAAACAGG	CTTTTTCTAG	TAATTATGAT	3840
TCCGGTGTTT	ATTCTTATTT	AACGCCTTAT	TTATCACACG	GTCGGTATTT	CAAACCATTA	3900
AATTTAGGTC	AGAAGATGAA	ATTAATAAAA	ATATATTTGA	AAAAGTTTTC	TCGCGTTCTT	3960
TGTCTTGCGA	TTGGATTGTC	ATCAGCATTT	ACATATAGTT	ATATAACCCA	ACCTAAGCCG	4020
GAGGTAAAA	AGGTAGTCTC	TCAGACCTAT	GATTTTGATA	AATTCACTAT	TGACTCTTCT	4080
CAGCGTCTTA	ATCTAAGCTA	TCGCTATGTT	TTCAAGGATT	CTAAGGGAAA	ATTAATTAAT	4140
AGCGACGATT	TACAGAAGCA	AGGTTATTCA	CTCACATATA	TTGATTTATG	TACTGTTTCC	4200
ATTAAAAAAG	GTAATTCAAA	TGAAATTGTT	AAATGTAATT	AATTTTGTTT	TCTTGATGTT	4260
TGTTTCATCA	TCTTCTTTTG	CTCAGGTAAT	TGAAATGAAT	AATTCGCCTC	TGCGCGATTT	4320
TGTAACCTGG	TATTCAAAGC	AATCAGGCGA	ATCCGTTATT	GTTTCTCCCG	ATGTAAGG	4380
TACTGTTACT	GATATTCAT	CTGACGTTAA	ACCTGAAAAT	CTACGCAATT	TCTTTATTTT	4440
TGTTTTACGT	GCTAATAATT	TTGATATGGT	TGGTTCAATT	CCTTCCATAA	TTCAGAAGTA	4500
TAATCCAAAC	AATCAGGATT	ATATTGATGA	ATTGCCATCA	TCTGATAATC	AGGAATATGA	4560
TGATAATTCC	GCTCCTTCTG	GTGGTTTCTT	TGTTCCGCAA	AATGATAATG	TACTCAAAC	4620
TTTTAAATTT	AATAACGTTT	GGGCAAAGGA	TTAATACGA	GTTGTCGAAT	TGTTTGTAAT	4680
GTCTAATACT	TCTAAATCCT	CAAATGTATT	ATCTATTGAC	GGCTCTAATC	TATTAGTTGT	4740
TAGTGACCT	AAAGATATTT	TAGATAACCT	TCCTCAATTC	CTTTCTACTG	TTGATTTGCC	4800
AACTGACCAG	ATATTGATTG	AGGGTTTGAT	ATTTGAGGTT	CAGCAAGGTG	ATGCTTTAGA	4860
TTTTTCATTT	GCTGCTGGCT	CTCAGCGTGG	CACTGTTGCA	GGCGGTGTTA	ATACTGACCG	4920
CCTCACCTCT	GTTTTATCTT	CTGCTGGTGG	TTCGTTCCGGT	ATTTTAAATG	GCGATGTTTT	4980
AGGGCTATCA	GTTTCGCGCAT	TAAAGACTAA	TAGCCATTCA	AAAATATTGT	CTGTGCCACG	5040



TATTCTTACG	CTTTCAGGTC	AGAAGGGTTC	TATCTCTGTT	GGCCAGAATG	TCCCTTTTAT	5100
TACTGGTCGT	GTGACTGGTG	AATCTGCCAA	TGTAAATAAT	CCATTTTCAGA	CGATTGAGCG	5160
TCAAAATGTA	GGTATTTCCA	TGAGCGTTTT	TCCTGTTGCA	ATGGCTGGCG	GTAATATTGT	5220
TCTGGATATT	ACCAGCAAGG	CCGATAGTTT	GAGTTCTTCT	ACTCAGGCAA	GTGATGTTAT	5280
TACTAATCAA	AGAAGTATTG	CTACAACGGT	TAATTTGCGT	GATGGACAGA	CTCTTTTACT	5340
CGGTGGCCTC	ACTGATTATA	AAAACACTTC	TCAAGATTCT	GGCGTACCGT	TCCTGTCTAA	5400
AATCCCTTTA	ATCGGCCTCC	TGTTTAGCTC	CCGCTCTGAT	TCCAACGAGG	AAAGCACGTT	5460
ATACGTGCTC	GTCAAAGCAA	CCATAGTACG	CGCCCTGTAG	CGGCGCATT	AGCGCGGCGG	5520
GTGTGGTGGT	TACGCGCAGC	GTGACCGCTA	CACTTGCCAG	CGCCCTAGCG	CCCGCTCCTT	5580
TCGCTTTCTT	CCCTTCCTTT	CTCGCCACGT	TCGCCGGCTT	TCCCCGTCAA	GCTCTAAATC	5640
GGGGGCTCCC	TTTAGGGTTC	CGATTTAGTG	CTTTACGGCA	CCTCGACCCC	AAAAAAGTTG	5700
ATTTGGGTGA	TGGTTCACGT	AGTGGGCCAT	CGCCCTGATA	GACGGTTTTT	CGCCCTTTGA	5760
CGTTGGAGTC	CACGTTCTTT	AATAGTGGAC	TCTTGTCCA	AACTGGAACA	ACACTCAACC	5820
CTATCTCGGG	CTATTCTTTT	GATTTATAAG	GGATTTTGCC	GATTTTCGAA	CCACCATCAA	5880
ACAGGATTTT	CGCCTGCTGG	GGCAAACCAG	CGTGGACCGC	TTGCTGCAAC	TCTCTCAGGG	5940
CCAGGCGGTG	AAGGGCAATC	AGCTGTTGCC	CGTCTCGCTG	GTGAAAAGAA	AAACCACCCT	6000
GGCGCCCAAT	ACGCAAACCG	CCTCTCCCCG	CGCGTTGGCC	GATTCATTAA	TGCAGCTGGC	6060
ACGACAGGTT	TCCCGACTGG	AAAGCGGGCA	GTGAGCGCAA	CGCAATTAAT	GTGAGTTAGC	6120
TCACTCATT	GGCACCCCAG	GCTTTACACT	TTATGCTTCC	GGCTCGTATG	TTGTGTGGAA	6180
TTGTGAGCGG	ATAACAATTT	CACACGCCAA	GGAGACAGTC	ATAATGAAAT	ACCTATTGCC	6240
TACGGCAGCC	GCTGGATTGT	TATTACTCGC	TGCCCCAACCA	GCCATGGCCG	AGCTCGTGAT	6300
GACCCAGACT	CCAGAATTCC	ATCCGGAATG	AGTGTTAATT	CTAGAACGCG	TAAGCTTGGC	6360
ACTGGCCGTC	GTTTTACAAC	GTCGTGACTG	GGAAAACCCCT	GGCGTTACCC	AACTTAATCG	6420
CCTTGACGCA	CACCCCCCTT	TCGCCAGCTG	GCGTAATAGC	GAAGAGGCCC	GCACCGATCG	6480
CCCTTCCCAA	CAGTTGCGCA	GCCTGAATGG	CGAATGGCGC	TTTGCTGGT	TTCCGGCACC	6540
AGAAGCGGTG	CCGGAAAGCT	GGCTGGAGTG	CGATCTTCCT	GAGGCCGATA	CGGTCTGTCG	6600
CCCCTCAAAC	TGGCAGATGC	ACGGTTACGA	TGCGCCCATC	TACACCAACG	TAACCTATCC	6660
CATTACGGTC	AATCCGCCGT	TTGTTCCCAC	GGAGAATCCG	ACGGGTTGTT	ACTCGCTCAC	6720
ATTTAATGTT	GATGAAAGCT	GGCTACAGGA	AGGCCAGACG	CGAATTATTT	TTGATGGCGT	6780
TCCTATTGGT	TAAAAAATGA	GCTGATTTAA	CAAAAATTTA	ACGCGAATTT	TAACAAAATA	6840

TTAACGTTTA	CAATTTAAAT	ATTTGCTTAT	ACAATCTTCC	TGTTTTTGGG	GCTTTTCTGA	6900
TTATCAACCG	GGGTACATAT	GATTGACATG	CTAGTTTAC	GATTACCGTT	CATCGATTCT	6960
CTTGTTTGCT	CCAGACTCTC	AGGCAATGAC	CTGATAGCCT	TTGTAGATCT	CTCAAAAATA	7020
GCTACCCTCT	CCGGCATTAA	TTTATCAGCT	AGAACGGTTG	AATATCATAT	TGATGGTGAT	7080
TTGACTGTCT	CCGGCCTTTC	TCACCCTTTT	GAATCTTTAC	CTACACATTA	CTCAGGCATT	7140
GCATTTAAAA	TATATGAGGG	TTCTAAAAAT	TTTTATCCTT	GCGTTGAAAT	AAAGGCTTCT	7200
CCCGCAAAAG	TATTACAGGG	TCATAATGTT	TTTGGTACAA	CCGATTTAGC	TTTATGCTCT	7260
GAGGCTTTAT	TGCTTAATTT	TGCTAATTCT	TTGCCTTGCC	TGTATGATTT	ATTGGACGTT	7320

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7445 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AATGCTACTA	CTATTAGTAG	AATTGATGCC	ACCTTTTCAG	CTCGCGCCCC	AAATGAAAAT	60
ATAGCTAAAC	AGGTTATTGA	CCATTTGCGA	AATGTATCTA	ATGGTCAAAC	TAAATCTACT	120
CGTTCGCAGA	ATTGGGAATC	AACTGTTACA	TGGAATGAAA	CTTCCAGACA	CCGTACTTTA	180
GTTGCATATT	TAAAACATGT	TGAGCTACAG	CACCAGATTC	AGCAATTAAG	CTCTAAGCCA	240
TCTGCAAAAA	TGACCTCTTA	TCAAAAAGGAG	CAATTAAAGG	TACTCTCTAA	TCCTGACCTG	300
TTGGAGTTTG	CTTCCGGTCT	GGTTCGCTTT	GAAGCTCGAA	TTAAAACGCG	ATATTTGAAG	360
TCTTTCGGGC	TTCCTCTTAA	TCTTTTTGAT	GCAATCCGCT	TTGCTTCTGA	CTATAATAGT	420
CAGGGTAAAG	ACCTGATTTT	TGATTTATGG	TCATTCTCGT	TTTCTGAACT	GTTTAAAGCA	480
TTTGAGGGGG	ATTCAATGAA	TATTTATGAC	GATTCCGCAG	TATTGGACGC	TATCCAGTCT	540
AAACATTTTA	CTATTACCCC	CTCTGGCAAA	ACTTCTTTTG	CAAAAGCCTC	TCGCTATTTT	600
GGTTTTTATC	GTCGTCTGGT	AAACGAGGGT	TATGATAGTG	TTGCTCTTAC	TATGCCTCGT	660
AATTCCTTTT	GGCGTTATGT	ATCTGCATTA	GTTGAATGTG	GTATTCCTAA	ATCTCAACTG	720
ATGAATCTTT	CTACCTGTAA	TAATGTTGTT	CCGTTAGTTC	GTTTTATTAA	CGTAGATTTT	780
TCTTCCCAAC	GTCCTGACTG	GTATAATGAG	CCAGTTCTTA	AAATCGCATA	AGGTAATTCA	840
CAATGATTAA	AGTTGAAATT	AAACCATCTC	AAGCCCAATT	TACTACTCGT	TCTGGTGTTT	900
CTCGTCAGGG	CAAGCCTTAT	TCACTGAATG	AGCAGCTTTG	TTACGTTGAT	TTGGGTAATG	960

AATATCCGGT	TCTTGTC AAG	ATTACTCTTG	ATGAAGGTCA	GCCAGCCTAT	GCGCCTGGTC	1020
TGTACACCGT	TCATCTGTCC	TCTTTC AAG	TTGGTCAGTT	CGGTTCCCTT	ATGATTGACC	1080
GTCTGCGCCT	CGTTCCGGCT	AAGTAACATG	GAGCAGGTCTG	CGGATTTCGA	CACAATTTAT	1140
CAGGCGATGA	TACAAATCTC	CGTTGTACTT	TGTTTCGCGC	TTGGTATAAT	CGCTGGGGGT	1200
CAAAGATGAG	TGTTTTAGTG	TATTCTTTTCG	CCTCTTTTCGT	TTTAGGTTGG	TGCCTTCGTA	1260
GTGGCATTAC	GTATTTTACC	CGTTTAATGG	AAACTTCCTC	ATGAAAAAGT	CTTTAGTCCT	1320
CAAAGCCTCT	GTAGCCGTTG	CTACCCTCGT	TCCGATGCTG	TCTTTCGCTG	CTGAGGGTGA	1380
CGATCCCGCA	AAAGCGGCCT	TTAACTCCCT	GCAAGCCTCA	GCGACCGAAT	ATATCGGTTA	1440
TGCGTGGGCG	ATGGTTGTTG	TCATTGTCTG	CGCAACTATC	GGTATCAAGC	TGTTTAAGAA	1500
ATTCACCTCG	AAAGCAAGCT	GATAAACCGA	TACAAATAAA	GGCTCCTTTT	GGAGCCTTTT	1560
TTTTTGGAGA	TTTTCAACGT	GAAAAAATTA	TTATTCGCAA	TTCCTTTAGT	TGTTCCTTTC	1620
TATTCTCACT	CCGCTGAAAC	TGTTGAAAGT	TGTTTAGCAA	AACCCCATAC	AGAAAATTCA	1680
TTTACTAACG	TCTGGAAAGA	CGACAAAAC	TTAGATCGTT	ACGCTAACTA	TGAGGGTTGT	1740
CTGTGGAATG	CTACAGGCGT	TGTAGTTTGT	ACTGGTGACG	AACTCAGTG	TTACGGTACA	1800
TGGGTTCCTA	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGCGGT	1860
TCTGAGGGTG	GCGGTTCTGA	GGGTGGCGGT	ACTAAACCTC	CTGAGTACGG	TGATACACCT	1920
ATTCCGGGCT	ATACTTATAT	CAACCCTCTC	GACGGCACTT	ATCCGCCTGG	TACTGAGCAA	1980
AACCCCGCTA	ATCCTAATCC	TTCTCTTGAG	GAGTCTCAGC	CTCTTAATAC	TTTCATGTTT	2040
CAGAATAATA	GGTTCCGAAA	TAGGCAGGGG	GCATTAAC	TTTATACGGG	CACTGTTACT	2100
CAAGGCACTG	ACCCCGTTAA	AACTTATTAC	CAGTACACTC	CTGTATCATC	AAAAGCCATG	2160
TATGACGCTT	ACTGGAACGG	TAAATTCAGA	GACTGCGCTT	TCCATTCTGG	CTTTAATGAA	2220
GATCCATTCTG	TTTGTGAATA	TCAAGGCCAA	TCGTCTGACC	TGCCTCAACC	TCCTGTCAAT	2280
GCTGGCGGCG	GCTCTGGTGG	TGGTTCTGGT	GGCGGCTCTG	AGGGTGGTGG	CTCTGAGGGT	2340
GGCGGTTCTG	AGGGTGGCGG	CTCTGAGGGA	GGCGGTTCCG	GTGGTGGCTC	TGGTTCGGGT	2400
GATTTTGATT	ATGAAAAGAT	GGCAAACGCT	AATAAGGGGG	CTATGACCGA	AAATGCCGAT	2460
GAAAACGCGC	TACAGTCTGA	CGCTAAAGGC	AAACTTGATT	CTGTCGCTAC	TGATTACGGT	2520
GCTGCTATCG	ATGGTTTCAT	TGGTGACGTT	TCCGGCCTTG	CTAATGGTAA	TGGTGCTACT	2580
GGTGATTTTG	CTGGCTCTAA	TTCCCAAATG	GCTCAAGTCG	GTGACGGTGA	TAATTCACCT	2640
TTAATGAATA	ATTTCCGTCA	ATATTTACCT	TCCCTCCCTC	AATCGGTTGA	ATGTCGCCCT	2700
TTTGTCTTTA	GCGCTGGTAA	ACCATATGAA	TTTTCTATTG	ATTGTGACAA	AATAAACTTA	2760

TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG	2820
TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT	2880
TATTATTGCG TTTCCCTCGGT TTCCTTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC	2940
TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CTATTTTATT GTTTCTTGCT CTTATTATTG	3000
GGCTTAACTC AATTCTTGTT GGTATCTCT CTGATATTAG CGCTCAATTA CCCTCTGACT	3060
TTGTTTCAAGG TGTTTCAAGT ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TATGTTATTC	3120
TCTCTGTAAA GGCTGCTATT TTCATTTTTG ACGTTAAACA AAAAATCGTT TCTTATTGG	3180
ATTGGGATAA ATAATATGGC TGTTTATTTT GTAAGTGGCA AATTAGGCTC TGGAAAGACG	3240
CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT	3300
CTTGATTTAA GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAAC GCCTCGCGTT	3360
CTTAGAATAC CGGATAAGCC TTCTATATCT GATTTGCTTG CTATTGGGCG CGGTAATGAT	3420
TCCTACGATG AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTGCAGTAC TTGGTTTAA	3480
ACCCGTTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTTCT ACATGCTCGT	3540
AAATTAGGAT GGGATATTAT TTTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG	3600
CGTTCTGCAT TAGCTGAACA TGTTGTTTAT TGTCGTCGTC TGGACAGAAT TACTTTACCT	3660
TTTGTCGGTA CTTTATATTC TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT	3720
GTTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT	3780
ACTGGTAAGA ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTTCTAG TAATTATGAT	3840
TCCGGTGTTT ATTCTTATT AACGCCTTAT TTATCACAGG GTCGGTATTT CAAACCATTA	3900
AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTTGA AAAAGTTTTT ACGCGTTCTT	3960
TGTCTTGCGA TTGGATTTGC ATCAGCATTT ACATATAGTT ATATAACCCA ACCTAAGCCG	4020
GAGGTAAAA AGGTAGTCTC TCAGACCTAT GATTTTGATA AATTCATAT TGACTCTTCT	4080
CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT	4140
AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTTATG TACTGTTTCC	4200
ATTAAAAAAG GTAATTCAAA TGAAATTGTT AAATGTAATT AATTTTGTTT TCTTGATGTT	4260
TGTTTCATCA TCTTCTTTTG CTCAGGTAAT TGAAATGAAT AATTCGCCTC TCGCGGATTT	4320
TGTAAGTTGG TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTCTCCCG ATGTAAAAGG	4380
TACTGTTACT GTATATTCAT CTGACGTAA ACCTGAAAAT CTACGCAATT TCTTTATTTT	4440
TGTTTTACGT GCTAATAATT TTGATATGGT TGTTTCAATT CCTTCCATAA TTCAGAAGTA	4500
TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA	4560

TGATAATTCC	GCTCCTTCTG	GTGGTTTCTT	TGTTCCGCAA	AATGATAATG	TTACTCAAAC	4620
TTTTAAAATT	AATAACGTTT	GGGCAAAGGA	TTTAATACGA	GTTGTCGAAT	TGTTTGTAAG	4680
GTCTAATACT	TCTAAATCCT	CAAATGTATT	ATCTATTGAC	GGCTCTAATC	TATTAGTTGT	4740
TAGTGACCT	AAAGATATTT	TAGATAACCT	TCCTCAATTC	CTTTCTACTG	TTGATTTGCC	4800
AACTGACCAG	ATATTGATTG	AGGGTTTGAT	ATTTGAGGTT	CAGCAAGGTG	ATGCTTTAGA	4860
TTTTTCATTT	GCTGCTGGCT	CTCAGCGTGG	CACTGTTGCA	GGCGGTGTTA	ATACTGACCG	4920
CCTCACCTCT	GTTTTATCTT	CTGCTGGTGG	TTCGTTCTGG	ATTTTTAATG	GCGATGTTTT	4980
AGGGCTATCA	GTTTCGCGCAT	TAAAGACTAA	TAGCCATTCA	AAAATATTGT	CTGTGCCACG	5040
TATTCTTACG	CTTTCAGGTC	AGAAGGGTTC	TATCTCTGTT	GGCCAGAATG	TCCCTTTTAT	5100
TACTGGTCGT	GTGACTGGTG	AATCTGCCAA	TGTAAATAAT	CCATTTTCAGA	CGATTGAGCG	5160
TCAAAATGTA	GGTATTTCCA	TGAGCGTTTT	TCCTGTTGCA	ATGGCTGGCG	GTAATATTGT	5220
TCTGGATATT	ACCAGCAAGG	CCGATAGTTT	GAGTTCTTCT	ACTCAGGCAA	GTGATGTTAT	5280
TACTAATCAA	AGAAGTATTG	CTACAACGGT	TAATTTGCGT	GATGGACAGA	CTCTTTTACT	5340
CGGTGGCCTC	ACTGATTATA	AAAACACTTC	TCAAGATTCT	GGCGTACCGT	TCCTGTCTAA	5400
AATCCCTTTA	ATCGGCCTCC	TGTTTAGCTC	CCGCTCTGAT	TCCAACGAGG	AAAGCACGTT	5460
ATACGTGCTC	GTCAAAGCAA	CCATAGTACG	CGCCCTGTAG	CGGCGCATTG	AGCGCGGCGG	5520
GTGTGGTGGT	TACGCGCAGC	GTGACCGCTA	CACTTGCCAG	CGCCCTAGCG	CCCGCTCCTT	5580
TCGCTTTCTT	CCCTTCCTTT	CTCGCCACGT	TCGCCGGCTT	TCCCCGTCAA	GCTCTAAATC	5640
GGGGGCTCCC	TTTAGGGTTC	CGATTTAGTG	CTTTACGGCA	CCTCGACCC	AAAAAAGTTG	5700
ATTTGGGTGA	TGGTTCACGT	AGTGGGCCAT	CGCCCTGATA	GACGGTTTTT	CGCCCTTTGA	5760
CGTTGGAGTC	CACGTTCTTT	AATAGTGGAC	TCTTGTTCCA	AACTGGAACA	AACTCAACC	5820
CTATCTCGGG	CTATTCTTTT	GATTTATAAG	GGATTTTGCC	GATTTTCGGAA	CCACCAATCAA	5880
ACAGGATTTT	CGCCTGCTGG	GGCAAACCAG	CGTGGACCGC	TTGCTGCAAC	TCTCTCAGGG	5940
CCAGGCGGTG	AAGGGCAATC	AGCTGTTGCC	CGTCTCGCTG	GTGAAAAGAA	AAACCACCTT	6000
GGCGCCCAAT	ACGCAAACCG	CCTCTCCCCG	CGCGTTGGCC	GATTCATTAA	TGCAGCTGGC	6060
ACGACAGGTT	TCCCGACTGG	AAAGCGGGCA	GTGAGCGCAA	CGCAATTAAT	GTGAGTTAGC	6120
TCATCATTA	GGCACCCAG	GCTTTACACT	TTATGCTTCC	GGCTCGTATG	TTGTGTGGAA	6180
TTGTGAGCGG	ATAACAATTT	CACACGCGTC	ACTTGGCACT	GGCCGTCGTT	TTACAACGTC	6240
GTGACTGGGA	AAACCCTGGC	GTTACCCAAG	CTTTGTACAT	GGAGAAAATA	AAGTGAAACA	6300
AAGCACTATT	GCACTGGCAC	TCTTACCGTT	ACCGTTACTG	TTTACCCCTG	TGACAAAAGC	6360

CGCCCAGGTC CAGCTGCTCG AGTCAGGCCT ATTGTGCCCCA GGGGATTGTA CTAGTGGATC	6420
CTAGGCTGAA GGCGATGACC CTGCTAAGGC TGCATTCAAT AGTTTACAGG CAAGTGCTAC	6480
TGAGTACATT GGCTACGCTT GGGCTATGGT AGTAGTTATA GTTGGTGCTA CCATAGGGAT	6540
TAAATTATTC AAAAAGTTTA CGAGCAAGGC TTCTTAAGCA ATAGCGAAGA GGGCCGCACC	6600
GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT GGCGCTTTGC CTGGTTTCCG	6660
GCACCAGAAG CGGTGCCGGA AAGCTGGCTG GAGTGCGATC TTCCTGAGGC CGATACGGTC	6720
GTCGTCCCCT CAAACTGGCA GATGCACGGT TACGATGCGC CCATCTACAC CAACGTAACC	6780
TATCCCATTA CGGTCAATCC GCCGTTTGTT CCCACGGAGA ATCCGACGGG TTGTTACTCG	6840
CTCACATTTA ATGTTGATGA AAGCTGGCTA CAGGAAGGCC AGACGCGAAT TATTTTTGAT	6900
GGCGTTCCTA TTGGTTAAAAA AATGAGCTGA TTTAACAAAA ATTTAACGCG AATTTTAACA	6960
AAATATTAAC GTTTACAATT TAAATATTTG CTTATACAAT CTCCTGTTT TTGGGGCTTT	7020
TCTGATTATC AACCAGGGTA CATATGATTG ACATGCTAGT TTTACGATTA CCGTTCATCG	7080
ATTCTCTTGT TTGCTCCAGA CTCTCAGGCA ATGACCTGAT AGCCTTTGTA GATCTCTCAA	7140
AAATAGCTAC CCTCTCCGGC ATTAATTTAT CAGCTAGAAC GGTTGAATAT CATATTGATG	7200
GTGATTTGAC TGTCTCCGGC CTTTCTCACC CTTTGAATC TTTACCTACA CATTACTCAG	7260
GCATTGCATT TAAAATATAT GAGGGTTCTA AAAATTTTTA TCCTTGCGTT GAAATAAAGG	7320
CTTCTCCCGC AAAAGTATTA CAGGGTCATA ATGTTTTTGG TACAACCGAT TTAGCTTTAT	7380
GCTCTGAGGC TTTATTGCTT AATTTTGCTA ATTCTTTGCC TTGCCTGTAT GATTTATTGG	7440
ACGTT	7445

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7409 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA	180
GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA	240
TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300

TTGGAGTTTG	CTTCCGGTCT	GGTTCGCTTT	GAAGCTCGAA	TTAAAACGCG	ATATTTGAAG	360
TCTTTCGGGC	TTCCTCTTAA	TCTTTTGTAT	GCAATCCGCT	TTGCTTCTGA	CTATAATAGT	420
CAGGGTAAAG	ACCTGATTTT	TGATTTATGG	TCATTCTCGT	TTTCTGAACT	GTTTAAAGCA	480
TTTGAGGGGG	ATTCAATGAA	TATTTATGAC	GATTCCGCAG	TATTGGACGC	TATCCAGTCT	540
AAACATTTTA	CTATTACCCC	CTCTGGCAAA	ACTTCTTTTG	CAAAAGCCTC	TCGCTATTTT	600
GGTTTTTATC	GTCGTCTGGT	AAACGAGGGT	TATGATAGTG	TTGCTCTTAC	TATGCCTCGT	660
AATTCCTTTT	GGCGTTATGT	ATCTGCATTA	GTTGAATGTG	GTATTCCTAA	ATCTCAACTG	720
ATGAATCTTT	CTACCTGTAA	TAATGTTGTT	CCGTTAGTTC	GTTTTATTAA	CGTAGATTTT	780
TCTTCCCAAC	GTCCTGACTG	GTATAATGAG	CCAGTTCTTA	AAATCGCATA	AGGTAATTCA	840
CAATGATTAA	AGTTGAAATT	AAACCATCTC	AAGCCCAATT	TACTACTCGT	TCTGGTGTTT	900
CTCGTCAGGG	CAAGCCTTAT	TCACTGAATG	AGCAGCTTTG	TTACGTTGAT	TTGGGTAATG	960
AATATCCGGT	TCTTGTC AAG	ATTACTCTTG	ATGAAGGTCA	GCCAGCCTAT	GCGCCTGGTC	1020
TGTACACCGT	TCATCTGTCC	TCTTTC AAG	TTGGTCAGTT	CGGTTCCCTT	ATGATTGACC	1080
GTCTGCGCCT	CGTTCCGGCT	AAGTAACATG	GAGCAGGTCG	CGGATTTCGA	CACAATTTAT	1140
CAGGCGATGA	TACAAATCTC	CGTTGTACTT	TGTTTCGCGC	TTGGTATAAT	CGCTGGGGGT	1200
CAAAGATGAG	TGTTTTAGTG	TATTCTTTTCG	CCTCTTTTCGT	TTTAGGTTGG	TGCCTTCGTA	1260
GTGGCATTAC	GTATTTTACC	CGTTTAATGG	AAACTTCCTC	ATGAAAAAGT	CTTTAGTCCT	1320
CAAAGCCTCT	GTAGCCGTTG	CTACCCCTCGT	TCCGATGCTG	TCTTTCGCTG	CTGAGGGTGA	1380
CGATCCCGCA	AAAGCGGCCT	TTAACTCCCT	GCAAGCCTCA	GCGACCGAAT	ATATCGGTTA	1440
TGCGTG GCG	ATGGTTGTTG	TCATTGTCGG	CGCAACTATC	GGTATCAAGC	TGTTTAAGAA	1500
ATTCACCTCG	AAAGCAAGCT	GATAAACC GA	TACAATTAAA	GGCTCCTTTT	GGAGCCTTTT	1560
TTTTTGGAGA	TTTTCAACGT	GAAAAAATTA	TTATTCGCAA	TTCTTTAGT	TGTTCTTTTC	1620
TATTCTCACT	CCGCTGAAAC	TGTTGAAAGT	TGTTTAGCAA	AACCCCATAC	AGAAAATTCA	1680
TTTACTAACG	TCTGGAAAGA	CGACAAAAC T	TTAGATCGTT	ACGCTAACTA	TGAGGGTTGT	1740
CTGTGGAATG	CTACAGGCGT	TGTAGTTTGT	ACTGGTGACG	AAACTCAGTG	TTACGGTACA	1800
TGGGTTCCTA	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGCGGT	1860
TCTGAGGGTG	GCGGTTCTGA	GGGTGGCGGT	ACTAAACCTC	CTGAGTACGG	TGATACACCT	1920
ATTCCGGGCT	ATACTTATAT	CAACCCCTCTC	GACGGCACTT	ATCCGCCTGG	TACTGAGCAA	1980
AACCCCGCTA	ATCCTAATCC	TTCTCTTGAG	GAGTCTCAGC	CTCTTAATAC	TTTCATGTTT	2040
CAGAATAATA	GGTTCCGAAA	TAGGCAGGGG	GCATTAACTG	TTTATACGGG	CACTGTTACT	2100

CAAGGCACTG	ACCCCGTTAA	AACTTATTAC	CAGTACACTC	CTGTATCATC	AAAAGCCATG	2160
TATGACGCTT	ACTGGAACGG	TAAATTCAGA	GACTGCGCTT	TCCATTCTGG	CTTTAATGAA	2220
GATCCATTTCG	TTTGTGAATA	TCAAGGCCAA	TCGTCTGACC	TGCCTCAACC	TCCTGTCAAT	2280
GCTGGCGGCG	GCTCTGGTGG	TGGTTCTGGT	GGCGGCTCTG	AGGGTGGTGG	CTCTGAGGGT	2340
GGCGGTTCTG	AGGGTGGCGG	CTCTGAGGGA	GGCGGTTCCG	GTGGTGGCTC	TGGTTCCGGT	2400
GATTTTGATT	ATGAAAAGAT	GGCAAACGCT	AATAAGGGGG	CTATGACCGA	AAATGCCGAT	2460
GAAAACGCGC	TACAGTCTGA	CGCTAAAGGC	AAACTTGATT	CTGTCGCTAC	TGATTACGGT	2520
GCTGCTATCG	ATGGTTTCAT	TGGTGACGTT	TCCGGCCTTG	CTAATGGTAA	TGGTGCTACT	2580
GGTGATTTTG	CTGGCTCTAA	TTCCCAAATG	GCTCAAGTCG	GTGACGGTGA	TAATTCACCT	2640
TTAATGAATA	ATTTCCGTCA	ATATTTACCT	TCCCTCCCTC	AATCGGTTGA	ATGTCGCCCT	2700
TTTGCTTTTA	GCGCTGGTAA	ACCATATGAA	TTTTCTATTG	ATTGTGACAA	AATAAACTTA	2760
TTCCGTGGTG	TCTTTGCGTT	TCTTTTATAT	GTTGCCACCT	TTATGTATGT	ATTTTCTACG	2820
TTTGCTAACA	TACTGCGTAA	TAAGGAGTCT	TAATCATGCC	AGTTCTTTTG	GGTATTCCGT	2880
TATTATTGCG	TTCCCTCGGT	TTCTTCTGG	TAACTTTGTT	CGGCTATCTG	CTTACTTTTC	2940
TTAAAAAGGG	CTTCGGTAAG	ATAGCTATTG	CTATTTCAAT	GTTTCTTGCT	CTTATTATTG	3000
GGCTTAACTC	AATTCTTG TG	GGTTATCTCT	CTGATATTAG	CGCTCAATTA	CCCTCTGACT	3060
TTGTTCAGGG	TGTTCAGTTA	ATTCTCCCGT	CTAATGCGCT	TCCCTGTTTT	TATGTTATTC	3120
TCTCTGTAAA	GGCTGCTATT	TTCATTTTTG	ACGTTAAACA	AAAAATCGTT	TCTTATTTGG	3180
ATTGGGATAA	ATAATATGGC	TGTTTATTTT	GTAACGGCA	AATTAGGCTC	TGGAAAGACG	3240
CTCGTTAGCG	TTGGTAAGAT	TCAGGATAAA	ATTGTAGCTG	GGTGCAAAAT	AGCAACTAAT	3300
CTTGATTTAA	GGCTTCAAAA	CCTCCCGCAA	GTCGGGAGGT	TCGCTAAAAC	GCCTCGCGTT	3360
CTTAGAATAC	CGGATAAGCC	TTCTATATCT	GATTTGCTTG	CTATTGGGCG	CGGTAATGAT	3420
TCCTACGATG	AAAATAAAAA	CGGCTTGCTT	GTTCTCGATG	AGTGCGGTAC	TTGGTTTAAT	3480
ACCCGTTCCT	GGAATGATAA	GGAAAGACAG	CCGATTATTG	ATTGGTTTCT	ACATGCTCGT	3540
AAATTAGGAT	GGGATATTAT	TTTTCTTGTT	CAGGACTTAT	CTATTGTTGA	TAAACAGGCG	3600
CGTTCTGCAT	TAGCTGAACA	TGTTGTTTAT	TGTCGTCGTC	TGGACAGAAT	TACTTTACCT	3660
TTTGTGCGTA	CTTTATATTC	TCTTATTACT	GGCTCGAAAA	TGCCTCTGCC	TAAATTACAT	3720
GTTGGCGTTG	TTAAATATGG	CGATTCTCAA	TTAAGCCCTA	CTGTTGAGCG	TTGGCTTTAT	3780
ACTGGTAAGA	ATTTGTATAA	CGCATATGAT	ACTAAACAGG	CTTTTCTAG	TAATTATGAT	3840
TCCGGTGTTT	ATTCTTATTT	AACGCCTTAT	TTATCACACG	GTCGGTATTT	CAAACCATTA	3900



AATTTAGGTC	AGAAGATGAA	GCTTACTAAA	ATATATTTGA	AAAAGTTTTTC	ACGCGTTCTT	3960
TGTCTTGCGA	TTGGATTTGC	ATCAGCATTT	ACATATAGTT	ATATAACCCA	ACCTAAGCCG	4020
GAGGTTAAAA	AGGTAGTCTC	TCAGACCTAT	GATTTTGATA	AATTCACATAT	TGACTCTTCT	4080
CAGCGTCTTA	ATCTAAGCTA	TCGCTATGTT	TTCAAGGATT	CTAAGGGAAA	ATTAATTAAT	4140
AGCGACGATT	TACAGAAGCA	AGGTTATTCA	CTCACATATA	TTGATTTATG	TACTGTTTCC	4200
ATTAAAAAAG	GTAATTCAAA	TGAAATTGTT	AAATGTAATT	AATTTTGTTT	TCTTGATGTT	4260
TGTTTCATCA	TCTTCTTTTG	CTCAGGTAAT	TGAAATGAAT	AATTCGCCTC	TGCGCGATTT	4320
TGTAACCTGG	TATTCAAAGC	AATCAGGCGA	ATCCGTTATT	GTTTCTCCCG	ATGTAAAAGG	4380
TACTGTTACT	GTATATTCAT	CTGACGTTAA	ACCTGAAAAT	CTACGCAATT	TCTTTATTTT	4440
TGTTTTACGT	GCTAATAATT	TTGATATGGT	TGGTTCAATT	CCTTCCATAA	TTCAGAAGTA	4500
TAATCCAAAC	AATCAGGATT	ATATTGATGA	ATTGCCATCA	TCTGATAATC	AGGAATATGA	4560
TGATAATTCC	GCTCCTTCTG	GTGGTTTCTT	TGTTCCGCAA	AATGATAATG	TTACTCAAAC	4620
TTTTAAAAAT	AATAACGTTC	GGGCAAAGGA	TTAATACGA	GTTGTCGAAT	TGTTTGTAAG	4680
GTCTAATACT	TCTAAATCCT	CAAATGTATT	ATCTATTGAC	GGCTCTAATC	TATTAGTTGT	4740
TAGTGCACCT	AAAGATATTT	TAGATAACCT	TCCTCAATTC	CTTTCTACTG	TTGATTTGCC	4800
AACTGACCAG	ATATTGATTG	AGGGTTTGAT	ATTTGAGGTT	CAGCAAGGTG	ATGCTTTAGA	4860
TTTTTCATTT	GCTGCTGGCT	CTCAGCGTGG	CACTGTTGCA	GGCGGTGTTA	ATACTGACCG	4920
CCTCACCTCT	GTTTTATCTT	CTGCTGGTGG	TTCGTTCCGT	ATTTTAAATG	GCGATGTTTT	4980
AGGGCTATCA	GTTCGCGCAT	TAAAGACTAA	TAGCCATTCA	AAAATATTGT	CTGTGCCACG	5040
TATTCTTACG	CTTTCAGGTC	AGAAGGGTTC	TATCTCTGTT	GGCCAGAATG	TCCCTTTTAT	5100
TACTGGTCGT	GTGACTGGTG	AATCTGCCAA	TGTAAATAAT	CCATTTCAGA	CGATTGAGCG	5160
TCAAAATGTA	GGTATTTCCA	TGAGCGTTTT	TCCTGTTGCA	ATGGCTGGCG	GTAATATTGT	5220
TCTGGATATT	ACCAGCAAGG	CCGATAGTTT	GAGTTCTTCT	ACTCAGGCAA	GTGATGTTAT	5280
TACTAATCAA	AGAAGTATTG	CTACAACGGT	TAATTTGCGT	GATGGACAGA	CTCTTTTACT	5340
CGGTGGCCTC	ACTGATTATA	AAAACACTTC	TCAAGATTCT	GGCGTACCGT	TCCTGTCTAA	5400
AATCCCTTTA	ATCGGCCTCC	TGTTTAGCTC	CCGCTCTGAT	TCCAACGAGG	AAAGCACGTT	5460
ATACGTGCTC	GTCAAAGCAA	CCATAGTACG	CGCCCTGTAG	CGGCGCATT	AGCGCGGCGG	5520
GTGTGGTGGT	TACGCGCAGC	GTGACCGCTA	CACTTGCCAG	CGCCCTAGCG	CCCCTCCTT	5580
TCGCTTTCTT	CCCTTCCTTT	CTCGCCACGT	TCGCCGGCTT	TCCCCGTCAA	GCTCTAAATC	5640
GGGGGCTCCC	TTTAGGGTTC	CGATTTAGTG	CTTTACGGCA	CCTCGACCCC	AAAAAATTG	5700

ATTTGGGTGA	TGGTTCACGT	AGTGGGCCAT	CGCCCTGATA	GACGGTTTTT	CGCCCTTTGA	5760
CGTTGGAGTC	CACGTTCTTT	AATAGTGGAC	TCTTGTTCCT	AACTGGAACA	AACTCAACC	5820
CTATCTCGGG	CTATTCTTTT	GATTTATAAG	GGATTTTGCC	GATTTTCGGAA	CCACCATCAA	5880
ACAGGATTTT	CGCCTGCTGG	GGCAAACCAG	CGTGGACCGC	TTGCTGCAAC	TCTCTCAGGG	5940
CCAGGCGGTG	AAGGGCAATC	AGCTGTTGCC	CGTCTCGCTG	GTGAAAAGAA	AAACCACCCT	6000
GGCGCCCAAT	ACGCAAACCG	CCTCTCCCCG	CGCGTTGGCC	GATTCATTAA	TGCAGCTGGC	6060
ACGACAGGTT	TCCCGACTGG	AAAGCGGGCA	GTGAGCGCAA	CGCAATTAAT	GTGAGTTAGC	6120
TCACTCATTA	GGCACCCAG	GCTTTACACT	TTATGCTTCC	GGCTCGTATG	TTGTGTGGAA	6180
TTGTGAGCGG	ATAACAATTT	CACACGCGTC	ACTTGGCACT	GGCCGTCGTT	TTACAACGTC	6240
GTGACTGGGA	AAACCCTGGC	GTTACCCAAG	CTTTGTACAT	GGAGAAAATA	AAGTGAAACA	6300
AAGCACTATT	GCACTGGCAC	TCTTACCGTT	ACTGTTTACC	CCTGTGGCAA	AAGCCTATGG	6360
GGGGTTTATG	ACTTCTGAGG	GATCCGGAGC	TGAAGGCGAT	GACCCTGCTA	AGGCTGCATT	6420
CAATAGTTTA	CAGGCAAGTG	CTACTGAGTA	CATTGGCTAC	GCTTGGGCTA	TGGTAGTAGT	6480
TATAGTTGGT	GCTACCATAG	GGATTAAATT	ATTCAAAAAG	TTTACGAGCA	AGGCTTCTTA	6540
AGCAATAGCG	AAGAGGCCCG	CACCGATCGC	CCTTCCCAAC	AGTTGCGCAG	CCTGAATGGC	6600
GAATGGCGCT	TTGCCTGGTT	TCCGGCACCA	GAAGCGGTGC	CGGAAAGCTG	GCTGGAGTGC	6660
GATCTTCCTG	AGGCCGATAC	GGTCGTCGTC	CCCTCAAAC	GGCAGATGCA	CGGTTACGAT	6720
GCGCCCATCT	ACACCAACGT	AACCTATCCC	ATTACGGTCA	ATCCGCCGTT	TGTTCCACG	6780
GAGAATCCGA	CGGGTTGTTA	CTCGCTCACA	TTTAATGTTG	ATGAAAAGCTG	GCTACAGGAA	6840
GGCCAGACGC	GAATTATTTT	TGATGGCGTT	CCTATTGGTT	AAAAAATGAG	CTGATTTAAC	6900
AAAAATTTAA	CGCGAATTTT	AACAAAATAT	TAACGTTTAC	AATTTAAATA	TTTGCTTATA	6960
CAATCTTCCT	GTTTTTGGGG	CTTTTCTGAT	TATCAACCGG	GGTACATATG	ATTGACATGC	7020
TAGTTTTACG	ATTACCGTTC	ATCGATTCTC	TTGTTTGCTC	CAGACTCTCA	GGCAATGACC	7080
TGATAGCCTT	TGTAGATCTC	TCAAAAATAG	CTACCCTCTC	CGGCATTAAT	TTATCAGCTA	7140
GAACGGTTGA	ATATCATATT	GATGGTGATT	TGACTGTCTC	CGGCCTTTCT	CACCCTTTTG	7200
AATCTTTACC	TACACATTAC	TCAGGCATTG	CATTTAAAT	ATATGAGGGT	TCTAAAAATT	7260
TTTATCCTTG	CGTTGAAATA	AAGGCTTCTC	CCGAAAAGT	ATTACAGGGT	CATAATGTTT	7320
TTGGTACAAC	CGATTTAGCT	TTATGCTCTG	AGGCTTTATT	GCTTAATTTT	GCTAATTCTT	7380
TGCCTTGCCT	GTATGATTTA	TTGGACGTT				7409

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7294 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA	180
GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA	240
TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCCGGTCT GGTTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG	360
TCTTTCGGGC TTCCTCTTAA TCTTTTGTAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGCAAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT	600
GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT	660
AATTCCTTTT GCGGTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT	780
TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA	840
CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTT	900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG	960
AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC	1020
TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC	1080
GTCTGCGCCT CGTTCCGGCT AAGTAACATG GAGCAGGTCTG CGGATTTCGA CACAATTTAT	1140
CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT	1200
CAAAGATGAG TGTTTTAGTG TATTCTTTCG CCTCTTTCGT TTTAGGTTGG TGCCTTCGTA	1260
GTGGCATTAC GTATTTTACC CGTTTAATGG AAACCTCCTC ATGAAAAAGT CTTTAGTCCT	1320
CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA	1380
CGATCCCGCA AAAGCGGCCT TTAACCTCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA	1440

TGCGTGGGCG	ATGGTTGTTG	TCATTGTCGG	CGCAACTATC	GGTATCAAGC	TGTTTAAGAA	1500
ATTCACCTCG	AAAGCAAGCT	GATAAACCGA	TACAATTAAA	GGCTCCTTTT	GGAGCCTTTT	1560
TTTTTGGAGA	TTTTCAACGT	AAAAAAATTA	TTATTTCGCA	TTCCTTTAGT	TGTTCCCTTT	1620
TATTCTCACT	CCGCTGAAAC	TGTTGAAAGT	TGTTTAGCAA	AACCCCATAC	AGAAAATTCA	1680
TTTACTAACG	TCTGGAAAGA	CGACAAAAC	TTAGATCGTT	ACGCTAACTA	TGAGGGTTGT	1740
CTGTGGAATG	CTACAGGCGT	TGTAGTTTGT	ACTGGTGACG	AAACTCAGTG	TTACGGTACA	1800
TGGGTTCCCTA	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGCGGT	1860
TCTGAGGGTG	GCGGTTCTGA	GGGTGGCGGT	ACTAAACCTC	CTGAGTACGG	TGATACACCT	1920
ATTCCGGGCT	ATACTTATAT	CAACCCTCTC	GACGGCACTT	ATCCGCCTGG	TACTGAGCAA	1980
AACCCCGCTA	ATCCTAATCC	TTCTCTTGAG	GAGTCTCAGC	CTCTTAATAC	TTTCATGTTT	2040
CAGAATAATA	GGTTCGAAA	TAGGCAGGGG	GCATTAAC	TTTATACGGG	CACTGTTACT	2100
CAAGGCACTG	ACCCCGTTAA	AACTTATTAC	CAGTACACTC	CTGTATCATC	AAAAGCCATG	2160
TATGACGCTT	ACTGGAACGG	TAAATTCAGA	GACTGCGCTT	TCCATTCTGG	CTTTAATGAA	2220
GATCCATTCG	TTTGTGAATA	TCAAGGCCAA	TCGTCTGACC	TGCCTCAACC	TCCTGTCAAT	2280
GCTGGCGGCG	GCTCTGGTGG	TGGTTCTGGT	GGCGGCTCTG	AGGGTGGTGG	CTCTGAGGGT	2340
GGCGGTTCTG	AGGGTGGCGG	CTCTGAGGGA	GGCGGTTCCG	GTGGTGGCTC	TGGTTCCGGT	2400
GATTTTGATT	ATGAAAAGAT	GGCAAACGCT	AATAAGGGGG	CTATGACCGA	AAATGCCGAT	2460
GAAAACGCGC	TACAGTCTGA	CGCTAAAGGC	AAACTTGATT	CTGTCGCTAC	TGATTACGGT	2520
GCTGCTATCG	ATGGTTTCAT	TGGTGACGTT	TCCGGCCTTG	CTAATGGTAA	TGGTGCTACT	2580
GGTGATTTTG	CTGGCTCTAA	TTCCCAAATG	GCTCAAGTCG	GTGACGGTGA	TAATTCACCT	2640
TTAATGAATA	ATTTCCGTCA	ATATTTACCT	TCCCTCCCTC	AATCGGTTGA	ATGTCGCCCT	2700
TTTGTCTTTA	GCGCTGGTAA	ACCATATGAA	TTTTCTATTG	ATTGTGACAA	AATAAACTTA	2760
TTCCGTGGTG	TCTTTGCGTT	TCTTTTATAT	GTTGCCACCT	TTATGTATGT	ATTTTCTACG	2820
TTTGCTAACA	TACTGCGTAA	TAAGGAGTCT	TAATCATGCC	AGTTCTTTTG	GGTATTCCGT	2880
TATTATTGCG	TTTCCTCGGT	TTCTTCTGG	TAACTTTGTT	CGGCTATCTG	CTTACTTTTC	2940
TTAAAAAGGG	CTTCGGTAAG	ATAGCTATTG	CTATTTTATT	GTTTCTTGCT	CTTATTATTG	3000
GGCTTAACTC	AATTCTTG	GGTTATCTCT	CTGATATTAG	CGCTCAATTA	CCCTCTGACT	3060
TTGTTTCAAGG	TGTTTCAAGT	ATTCTCCCGT	CTAATGCGCT	TCCCTGTTTT	TATGTTATTTC	3120
TCTCTGTAAA	GGCTGCTATT	TTCATTTTTG	ACGTAAACA	AAAAATCGTT	TCTTATTG	3180
ATTGGGATAA	ATAATATGGC	TGTTTATTTT	GTAAGTGGCA	AATTAGGCTC	TGAAAGACG	3240

CTCGTTAGCG	TTGGTAAGAT	TCAGGATAAA	ATTGTAGCTG	GGTGCAAAAT	AGCAACTAAT	3300
CTTGATTTAA	GGCTTCAAAA	CCTCCCGCAA	GTCGGGAGGT	TCGCTAAAAC	GCCTCGCGTT	3360
CTTAGAATAC	CGGATAAGCC	TTCTATATCT	GATTTGCTTG	CTATTGGGCG	CGGTAATGAT	3420
TCCTACGATG	AAAATAAAAA	CGGCTTGCTT	GTTCTCGATG	AGTGCGGTAC	TTGGTTTAAT	3480
ACCCGTTCTT	GGAATGATAA	GGAAAGACAG	CCGATTATTG	ATTGGTTTCT	ACATGCTCGT	3540
AAATTAGGAT	GGGATATTAT	CTTCCTTGTT	CAGGACTTAT	CTATTGTTGA	TAAACAGGCG	3600
CGTTCTGCAT	TAGCTGAACA	TGTTGTTTAT	TGTCGTCGTC	TGGACAGAAT	TACTTTACCT	3660
TTTGTCGGTA	CTTTATATTC	TCTTATTACT	GGCTCGAAAA	TGCCTCTGCC	TAAATTACAT	3720
GTTGGCGTTG	TTAAATATGG	CGATTCTCAA	TTAAGCCCTA	CTGTTGAGCG	TTGGCTTTAT	3780
ACTGGTAAGA	ATTTGTATAA	CGCATATGAT	ACTAAACAGG	CTTTTTCTAG	TAATTATGAT	3840
TCCGGTGTTT	ATTCTTATTT	AACGCCTTAT	TTATCACACG	GTCGGTATTT	CAAACCATTA	3900
AATTTAGGTC	AGAAGATGAA	GCTTACTAAA	ATATATTTGA	AAAAGTTTTC	ACGCGTTCTT	3960
TGTCTTGCGA	TTGGATTGTC	ATCAGCATT	ACATATAGTT	ATATAACCCA	ACCTAAGCCG	4020
GAGGTAAAAA	AGGTAGTCTC	TCAGACCTAT	GATTTTGATA	AATTCACTAT	TGACTCTTCT	4080
CAGCGTCTTA	ATCTAAGCTA	TCGCTATGTT	TTCAAGGATT	CTAAGGGAAA	ATTAATTAAT	4140
AGCGACGATT	TACAGAAGCA	AGGTTATTCA	CTCACATATA	TTGATTTATG	TACTGTTTCC	4200
ATTAAAAAGG	TAATTCAAAT	GAAATTGTTA	AATGTAATTA	ATTTTGTTTT	CTTGATGTTT	4260
GTTTCATCAT	CTTCTTTTGC	TCAGGTAATT	GAAATGAATA	ATTCGCCTCT	GCGCGATTTT	4320
GTAACTTGGT	ATTCAAAGCA	ATCAGGCGAA	TCCGTTATTG	TTTCTCCCGA	TGTAAAAGGT	4380
ACTGTTACTG	TATATTCATC	TGACGTTAAA	CCTGAAAATC	TACGCAATTT	CTTTATTTCT	4440
GTTTTACGTG	CTAATAATTT	TGATATGGTT	GGTTCAATTC	CTTCCATTAT	TTAGAAGTAT	4500
AATCCAAACA	ATCAGGATTA	TATTGATGAA	TTGCCATCAT	CTGATAATCA	GGAATATGAT	4560
GATAATTCCG	CTCCTTCTGG	TGGTTTCTTT	GTTCCGCAAA	ATGATAATGT	TACTCAAAC	4620
TTTAAAATTA	ATAACGTTTCG	GGCAAAGGAT	TTAATACGAG	TTGTCGAATT	GTTTGTAAG	4680
TCTAATACTT	CTAAATCCTC	AAATGTATTA	TCTATTGACG	GCTCTAATCT	ATTAGTTGTT	4740
AGTGACCTTA	AAGATATTTT	AGATAACCTT	CCTCAATTCC	TTTCTACTGT	TGATTTGCCA	4800
ACTGACCAGA	TATTGATTGA	GGGTTTGATA	TTTGAGGTTT	AGCAAGGTGA	TGCTTTAGAT	4860
TTTTCATTTG	CTGCTGGCTC	TCAGCGTGGC	ACTGTTGCAG	GCGGTGTTAA	TACTGACCGC	4920
CTCACCTCTG	TTTTATCTTC	TGCTGGTGGT	TCGTTCCGTA	TTTTTAATGG	CGATGTTTTA	4980
GGGCTATCAG	TTCGCGCATT	AAAGACTAAT	AGCCATTCAA	AAATATTGTC	TGTGCCACGT	5040

ATTCTTACGC	TTTCAGGTCA	GAAGGGTTCT	ATCTCTGTTG	GCCAGAATGT	CCCTTTTATT	5100
ACTGGTCGTG	TGACTGGTGA	ATCTGCCAAT	GTAAATAATC	CATTTTCAGAC	GATTGAGCGT	5160
CAAAATGTAG	GTATTTCCAT	GAGCGTTTTT	CCTGTTGCAA	TGGCTGGCGG	TAATATTGTT	5220
CTGGATATTA	CCAGCAAGGC	CGATAGTTTG	AGTTCTTCTA	CTCAGGCAAG	TGATGTTATT	5280
ACTAATCAAA	GAAGTATTGC	TACAACGGTT	AATTTGCGTG	ATGGACAGAC	TCTTTTACTC	5340
GGTGGCCTCA	CTGATTATAA	AAACACTTCT	CAAGATTCTG	GCGTACCGTT	CCTGTCTAAA	5400
ATCCCTTTAA	TCGGCCTCCT	GTTTAGCTCC	CGCTCTGATT	CCAACGAGGA	AAGCACGTTA	5460
TACGTGCTCG	TCAAAGCAAC	CATAGTACGC	GCCCTGTAGC	GGCGCATTA	GCGCGGCGGG	5520
TGTGGTGGTT	ACGCGCAGCG	TGACCGCTAC	ACTTGCCAGC	GCCCTAGCGC	CCGCTCCTTT	5580
CGCTTTCTTC	CCTTCCTTTC	TCGCCACGTT	CGCCGGCTTT	CCCCGTCAAG	CTCTAAATCG	5640
GGGGCTCCCT	TTAGGGTTCC	GATTTAGTGC	TTTACGGCAC	CTCGACCCCA	AAAACTTGA	5700
TTTGGGTGAT	GGTTCACGTA	GTGGGCCATC	GCCCTGATAG	ACGGTTTTTC	GCCCTTTGAC	5760
GTTGGAGTCC	ACGTTCTTTA	ATAGTGGACT	CTTGTTCCAA	ACTGGAACAA	CACTCAACCC	5820
TATCTCGGGC	TATTCTTTTG	ATTTATAAGG	GATTTTGCCG	ATTCGGAAC	CACCATCAAA	5880
CAGGATTTTC	GCCTGCTGGG	GCAAACCAGC	GTGGACCGCT	TGCTGCAACT	CTCTCAGGGC	5940
CAGGCGGTGA	AGGGCAATCA	GCTGTTGCCC	GTCTCGCTGG	TGAAAAGAAA	AACCACCCTG	6000
GCGCCCAATA	CGCAAACCGC	CTCTCCCCGC	GCGTTGGCCG	ATTCATTAAT	GCAGCTGGCA	6060
CGACAGGTTT	CCCGACTGGA	AAGCGGGCAG	TGAGCGCAAC	GCAATTAATG	TGAGTTAGCT	6120
CACTCATTAG	GCACCCCAGG	CTTTACACTT	TATGCTTCCG	GCTCGTATGT	TGTGTGGAAT	6180
TGTGAGCGGA	TAACAATTTT	ACACAGGAAA	CAGCTATGAC	CAGGATGTAC	GAATTCGCAG	6240
GTAGGAGAGC	TCGGCGGATC	CGAGGCTGAA	GGCGATGACC	CTGCTAAGGC	TGCATTCAAT	6300
AGTTTACAGG	CAAGTGCTAC	TGAGTACATT	GGCTACGCTT	GGGCTATGGT	AGTAGTTATA	6360
GTTGGTGCTA	CCATAGGGAT	TAAATTATTC	AAAAAGTTTA	CGAGCAAGGC	TTCTTAACCA	6420
GCTGGCGTAA	TAGCGAAGAG	GCCCGCACC	ATCGCCCTTC	CCAACAGTTG	CGCAGCCTGA	6480
ATGGCGAATG	GCGCTTTGCC	TGGTTTCCGG	CACCAGAAGC	GGTGCCGGA	AGCTGGCTGG	6540
AGTGCGATCT	TCCTGAGGCC	GATACGGTCG	TCGTCCCCTC	AAACTGGCAG	ATGCACGGTT	6600
ACGATGCGCC	CATCTACACC	AACGTAACCT	ATCCCATTAC	GGTCAATCCG	CCGTTTGTTT	6660
CCACGGAGAA	TCCGACGGGT	TGTTACTCGC	TCACATTTAA	TGTTGATGAA	AGCTGGCTAC	6720
AGGAAGGCCA	GACGCGAATT	ATTTTTGATG	GCGTTCCTAT	TGGTTAAAAA	ATGAGCTGAT	6780
TTAACAAAAA	TTTAACGCGA	ATTTTAACAA	AATATTAACG	TTTACAATTT	AAATATTTGC	6840

TTATACAATC TTCCTGTTTT TGGGGCTTTT CTGATTATCA ACCGGGGTAC ATATGATTGA	6900
CATGCTAGTT TTACGATTAC CGTTCATCGA TTCTCTTGTT TGCTCCAGAC TCTCAGGCAA	6960
TGACCTGATA GCCTTTGTAG ATCTCTCAA AATAGCTACC CTCTCCGGCA TTAATTTATC	7020
AGCTAGAACG GTTGAATATC ATATTGATGG TGATTTGACT GTCTCCGGCC TTTCTCACCC	7080
TTTTGAATCT TTACCTACAC ATTACTCAGG CATTGCATTT AAAATATATG AGGGTTCTAA	7140
AAATTTTTAT CCTTGC GTT AAATAAAGGC TTCTCCCGCA AAAGTATTAC AGGGTCATAA	7200
TGTTTTTGGT ACAACCGATT TAGCTTTATG CTCTGAGGCT TTATTGCTTA ATTTTGCTAA	7260
TTCTTTGCCT TGCCTGTATG ATTTATTGGA CGTT	7294

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7394 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCAGACA CCGTACTTTA	180
GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA	240
TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCGGTCT GGTTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG	360
TCTTTCGGGC TTCCTCTTAA TCTTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGCAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT	600
GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT	660
AATTCCTTTT GGC GTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT	780
TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA	840
CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTT	900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG	960

AATATCCGGT	TCTTGTC AAG	ATTACTCTTG	ATGAAGGTCA	GCCAGCCTAT	GCGCCTGGTC	1020
TGTACACCGT	TCATCTGTCC	TCTTTCAAAG	TTGGTCAGTT	CGGTTCCCTT	ATGATTGACC	1080
GTCTGCGCCT	CGTTCCGGCT	AAGTAACATG	GAGCAGGTCTG	CGGATTTCTGA	CACAATTTAT	1140
CAGGCGATGA	TACAAATCTC	CGTTGTACTT	TGTTTCGCGC	TTGGTATAAT	CGCTGGGGGT	1200
CAAAGATGAG	TGTTTTAGTG	TATTCTTTTCG	CCTCTTTTCGT	TTTAGGTTGG	TGCCTTCGTA	1260
GTGGCATTAC	GTATTTTACC	CGTTTAATGG	AAACTTCCTC	ATGAAAAAGT	CTTTAGTCCT	1320
CAAAGCCTCT	GTAGCCGTTG	CTACCCCTCGT	TCCGATGCTG	TCTTTCGCTG	CTGAGGGTGA	1380
CGATCCCGCA	AAAGCGGCCT	TTAACTCCCT	GCAAGCCTCA	GCGACCGAAT	ATATCGGTTA	1440
TGCGTGGGCG	ATGGTTGTTG	TCATTGTCGG	CGCAACTATC	GGTATCAAGC	TGTTTAAGAA	1500
ATTCACCTCG	AAAGCAAGCT	GATAAACCGA	TACAATTAAA	GGCTCCTTTT	GGAGCCTTTT	1560
TTTTTGAGAG	TTTTCAACGT	GAAAAAATTA	TTATTCGCAA	TTCTTTTAGT	TGTTCTTTTC	1620
TATTCTCACT	CCGCTGAAAC	TGTTGAAAGT	TGTTTAGCAA	AACCCCATAC	AGAAAATTCA	1680
TTTACTAACG	TCTGGAAAGA	CGACAAAAC	TTAGATCGTT	ACGCTAACTA	TGAGGGTTGT	1740
CTGTGGAATG	CTACAGGCGT	TGTAGTTTGT	ACTGGTGACG	AAACTCAGTG	TTACGGTACA	1800
TGGGTTCCCTA	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGCGGT	1860
TCTGAGGGTG	GCGGTTCTGA	GGGTGGCGGT	ACTAAACCTC	CTGAGTACGG	TGATACACCT	1920
ATTCCGGGCT	ATACTTATAT	CAACCCTCTC	GACGGCACTT	ATCCGCCTGG	TACTGAGCAA	1980
AACCCCGCTA	ATCCTAATCC	TTCTCTTGAG	GAGTCTCAGC	CTCTTAATAC	TTTCATGTTT	2040
CAGAATAATA	GGTTCCGAAA	TAGGCAGGGG	GCATTAAGTG	TTTATACGGG	CACTGTTACT	2100
CAAGGCACTG	ACCCCGTTAA	AACTTATTAC	CAGTACACTC	CTGTATCATC	AAAAGCCATG	2160
TATGACGCTT	ACTGGAACGG	TAAATTCAGA	GACTGCGCTT	TCCATTCTGG	CTTTAATGAA	2220
GATCCATTCG	TTTGTGAATA	TCAAGGCCAA	TCGTCTGACC	TGCCTCAACC	TCCTGTCAAT	2280
GCTGGCGGCG	GCTCTGGTGG	TGGTTCTGGT	GGCGGCTCTG	AGGGTGGTGG	CTCTGAGGGT	2340
GGCGGTTCTG	AGGGTGGCGG	CTCTGAGGGA	GGCGGTTCCG	GTGGTGGCTC	TGGTTCCGGT	2400
GATTTTGATT	ATGAAAAGAT	GGCAAACGCT	AATAAGGGGG	CTATGACCGA	AAATGCCGAT	2460
GAAAACGCGC	TACAGTCTGA	CGCTAAAGGC	AAACTTGATT	CTGTCGCTAC	TGATTACGGT	2520
GCTGCTATCG	ATGGTTTCAT	TGGTGACGTT	TCCGGCCTTG	CTAATGGTAA	TGGTGCTACT	2580
GGTGATTTTG	CTGGCTCTAA	TTCCCAAATG	GCTCAAGTCG	GTGACGGTGA	TAATTCACCT	2640
TTAATGAATA	ATTTCCGTCA	ATATTTACCT	TCCCTCCCTC	AATCGGTTGA	ATGTCGCCCT	2700
TTTGTCTTTA	GCGCTGGTAA	ACCATATGAA	TTTTCTATTG	ATTGTGACAA	AATAAACTTA	2760



TTCCGTGGTG	TCTTTGCGTT	TCTTTTATAT	GTTGCCACCT	TTATGTATGT	ATTTTCTACG	2820
TTTGCTAACA	TACTGCGTAA	TAAGGAGTCT	TAATCATGCC	AGTTCTTTTG	GGTATTCGGT	2880
TATTATTGCG	TTTCCTCGGT	TTCCTTCTGG	TAACTTTGTT	CGGCTATCTG	CTTACTTTTC	2940
TTAAAAAGGG	CTTCGGTAAG	ATAGCTATTG	CTATTTCAAT	GTTTCTTGCT	CTTATTATTG	3000
GGCTTAACTC	AATTCTTGTTG	GGTTATCTCT	CTGATATTAG	CGCTCAATTA	CCCTCTGACT	3060
TTGTTCAAGG	TGTTCAAGTTA	ATTCTCCCGT	CTAATGCGCT	TCCCTGTTTT	TATGTTATTC	3120
TCTCTGTAAA	GGCTGCTATT	TTCATTTTTG	ACGTTAAACA	AAAAATCGTT	TCTTATTTGG	3180
ATTGGGATAA	ATAATATGGC	TGTTTATTTT	GTAAGTGGCA	AATTAGGCTC	TGGAAAGACG	3240
CTCGTTAGCG	TTGGTAAGAT	TTAGGATAAA	ATTGTAGCTG	GGTGCAAAAT	AGCAACTAAT	3300
CTTGATTTAA	GGCTTCAAAA	CCTCCCGCAA	GTCGGGAGGT	TCGCTAAAAC	GCCTCGCGTT	3360
CTTAGAATAC	CGGATAAGCC	TTCTATATCT	GATTTGCTTG	CTATTGGGCG	CGGTAATGAT	3420
TCCTACGATG	AAAATAAAAA	CGGCTTGCTT	GTTCTCGATG	AGTGCGGTAC	TTGGTTTAAT	3480
ACCCGTTCTT	GGAATGATAA	GGAAAGACAG	CCGATTATTG	ATTGGTTTCT	ACATGCTCGT	3540
AAATTAGGAT	GGGATATTAT	TTTTCTTGTT	CAGGACTTAT	CTATTGTTGA	TAAACAGGCG	3600
CGTTCTGCAT	TAGCTGAACA	TGTTGTTTAT	TGTCGTCGTC	TGGACAGAAT	TACTTTACCT	3660
TTTGTCGGTA	CTTTATATTC	TCTTATTACT	GGCTCGAAAA	TGCCTCTGCC	TAAATTACAT	3720
GTTGGCGTTG	TTAAATATGG	CGATTCTCAA	TTAAGCCCTA	CTGTTGAGCG	TTGGCTTTAT	3780
ACTGGTAAGA	ATTTGTATAA	CGCATATGAT	ACTAAACAGG	CTTTTCTAG	TAATTATGAT	3840
TCCGGTGTTT	ATTCTTATTT	AACGCCCTAT	TTATCACACG	GTCGGTATTT	CAAACCATTA	3900
AATTTAGGTC	AGAAGATGAA	GCTTACTAAA	ATATATTTGA	AAAAGTTTTC	ACGCGTTCTT	3960
TGTCTTGCGA	TTGGATTTGC	ATCAGCATTT	ACATATAGTT	ATATAACCCA	ACCTAAGCCG	4020
GAGGTAAAAA	AGGTAGTCTC	TCAGACCTAT	GATTTTGATA	AATTCATAT	TGACTCTTCT	4080
CAGCGTCTTA	ATCTAAGCTA	TCGCTATGTT	TTCAAGGATT	CTAAGGGAAA	ATTAATTAAT	4140
AGCGACGATT	TACAGAAGCA	AGGTTATTCA	CTCACATATA	TTGATTTATG	TACTGTTTCC	4200
ATTAAAAAAG	GTAATTCAAA	TGAAATTGTT	AAATGTAATT	AATTTTGTTT	TCTTGATGTT	4260
TGTTTCATCA	TCTTCTTTTG	CTCAGGTAAT	TGAAATGAAT	AATTCGCCTC	TGCGCGATTT	4320
TGTAACCTGG	TATTCAAAGC	AATCAGGCGA	ATCCGTTATT	GTTTCTCCCG	ATGTAAAAGG	4380
TACTGTTACT	GTATATTCAT	CTGACGTTAA	ACCTGAAAAT	CTACGCAATT	TCTTTATTTT	4440
TGTTTTACGT	GCTAATAATT	TTGATATGGT	TGGTTCAATT	CCTTCCATAA	TTCAGAAGTA	4500
TAATCCAAAC	AATCAGGATT	ATATTGATGA	ATTGCCATCA	TCTGATAATC	AGGAATATGA	4560

TGATAATTCC	GCTCCTTCTG	GTGGTTTCTT	TGTTCCGCAA	AATGATAATG	TTACTCAAAC	4620
TTTTAAAATT	AATAACGTTC	GGGCAAAGGA	TTAATACGA	GTTGTCGAAT	TGTTTGTAAG	4680
GTCTAATACT	TCTAAATCCT	CAAATGTATT	ATCTATTGAC	GGCTCTAATC	TATTAGTTGT	4740
TAGTGCACCT	AAAGATATTT	TAGATAACCT	TCCTCAATTC	CTTTCTACTG	TTGATTTGCC	4800
AACTGACCAG	ATATTGATTG	AGGGTTTGAT	ATTTGAGGTT	CAGCAAGGTG	ATGCTTTAGA	4860
TTTTTCATTT	GCTGCTGGCT	CTCAGCGTGG	CACTGTTGCA	GGCGGTGTTA	ATACTGACCG	4920
CCTCACCTCT	GTTTTATCTT	CTGCTGGTGG	TTCGTTCTGGT	ATTTTAAATG	GCGATGTTTT	4980
AGGGCTATCA	GTTGCGCAT	TAAAGACTAA	TAGCCATTCA	AAAATATTGT	CTGTGCCACG	5040
TATTCTTACG	CTTTCAGGTC	AGAAGGGTTC	TATCTCTGTT	GGCCAGAATG	TCCCTTTTAT	5100
TACTGGTCGT	GTGACTGGTG	AATCTGCCAA	TGTAAATAAT	CCATTTTCTA	CGATTGAGCG	5160
TCAAAATGTA	GGTATTTCCA	TGAGCGTTTT	TCCTGTTGCA	ATGGCTGGCG	GTAATATTGT	5220
TCTGGATATT	ACCAGCAAGG	CCGATAGTTT	GAGTTCTTCT	ACTCAGGCAA	GTGATGTTAT	5280
TACTAATCAA	AGAAGTATTG	CTACAACGGT	TAATTTGCGT	GATGGACAGA	CTCTTTTACT	5340
CGGTGGCCTC	ACTGATTATA	AAAACACTTC	TCAAGATTCT	GGCGTACCGT	TCCTGTCTAA	5400
AATCCCTTTA	ATCGGCCTCC	TGTTTAGCTC	CCGCTCTGAT	TCCAACGAGG	AAAGCACGTT	5460
ATACGTGCTC	GTCAAAGCAA	CCATAGTACG	CGCCCTGTAG	CGGCGCATT	AGCGCGGCGG	5520
GTGTGGTGGT	TACGCGCAGC	GTGACCGCTA	CACTTGCCAG	CGCCCTAGCG	CCCGCTCCTT	5580
TCGCTTTCTT	CCCTTCCTTT	CTCGCCACGT	TCGCCGGCTT	TCCCCGTCAA	GCTCTAAATC	5640
GGGGGCTCCC	TTTAGGGTTC	CGATTTAGTG	CTTTACGGCA	CCTCGACCCC	AAAAAACTTG	5700
ATTTGGGTGA	TGGTTCACGT	AGTGGGCCAT	CGCCCTGATA	GACGGTTTTT	CGCCCTTTGA	5760
CGTTGGAGTC	CACGTTCTTT	AATAGTGGAC	TCTTGTTCCA	AACTGGAACA	AACTCAACC	5820
CTATCTCGGG	CTATTCTTTT	GATTTATAAG	GGATTTTGCC	GATTTGCGAA	CCACCATCAA	5880
ACAGGATTTT	CGCCTGCTGG	GGCAAACCAG	CGTGGACCGC	TTGCTGCAAC	TCTCTCAGGG	5940
CCAGGCGGTG	AAGGGCAATC	AGCTGTTGCC	CGTCTCGCTG	GTGAAAAGAA	AAACCACCTT	6000
GGCGCCCAAT	ACGCAAACCG	CCTCTCCCCG	CGCGTTGGCC	GATTCATTAA	TGCAGCTGGC	6060
ACGACAGGTT	TCCCGACTGG	AAAGCGGGCA	GTGAGCGCAA	CGCAATTAAT	GTGAGTTAGC	6120
TCACTCATT	GGCACCCAG	GCTTTACACT	TTATGCTTCC	GGCTCGTATG	TTGTGTGGAA	6180
TTGTGAGCGG	ATAACAATTT	CACACGCGTC	ACTTGGCACT	GGCCGTCGTT	TTACAACGTC	6240
GTGACTGGGA	AAACCCTGGC	GTTACCCAAG	CTTTGTACAT	GGAGAAAATA	AAGTGAAACA	6300
AAGCACTATT	GCACTGGCAC	TCTTACCGTT	ACTGTTTACC	CCTGTGGCAA	AAGCCCTTCT	6360

GAGGCATCCG GGAGCTGAAG GCGATGACCC TGCTAAGGCT GCATTCAATA GTTTACAGGC	6420
AAGTGCTACT GAGTACATTG GCTACGCTTG GGCTATGGTA GTAGTTATAG TTGGTGCTAC	6480
CATAGGGATT AAATTATTCA AAAAGTTTAC GAGCAAGGCT TCTTAAGCAA TAGCGAAGAG	6540
CCCCGCACCG ATCGCCCTTC CCAACAGTTG CGCAGCCTGA ATGGCGAATG GCGCTTTGCC	6600
TGGTTTCCGG CACCAGAAGC GGTGCCGGA AGCTGGCTGG AGTGCGATCT TCCTGAGGCC	6660
GATACGGTCG TCGTCCCCTC AAAC TGGCAG ATGCACGGTT ACGATGCGCC CATCTACACC	6720
AACGTAACCT ATCCCATTAC GGTCAATCCG CCGTTTGTTT CCACGGAGAA TCCGACGGGT	6780
TGTTACTCGC TCACATTTAA TGTTGATGAA AGCTGGCTAC AGGAAGGCCA GACGCGAATT	6840
ATTTTTGATG GCGTTCCTAT TGGTTAAAAA ATGAGCTGAT TTAACAAAAA TTTAACGCGA	6900
ATTTTAACAA AATATTAACG TTTACAATTT AAATATTTGC TTATACAATC TTCCTGTTTT	6960
TGGGGCTTTT CTGATTATCA ACCGGGGTAC ATATGATTGA CATGCTAGTT TTACGATTAC	7020
CGTTCATCGA TTCTCTTGTT TGCTCCAGAC TCTCAGGCAA TGACCTGATA GCCTTTGTAG	7080
ATCTCTCAAA AATAGCTACC CTCTCCGGCA TTAATTTATC AGCTAGAACG GTTGAATATC	7140
ATATTGATGG TGATTTGACT GTCTCCGGCC TTTCTCACC TTTTGAATCT TTACCTACAC	7200
ATTACTCAGG CATTGCATTT AAAATATATG AGGGTTCTAA AAATTTTAT CCTTGCCTTG	7260
AAATAAAGGC TTCTCCCGCA AAAGTATTAC AGGGTCATAA TGTTTTTGGT ACAACCGATT	7320
TAGCTTTATG CTCTGAGGCT TTATTGCTTA ATTTTGCTAA TTCTTTGCCT TGCCTGTATG	7380
ATTTATTGGA CGTT	7394

## (2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 37 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GATCCTAGGC TGAAGGCGAT GACCCTGCTA AGGCTGC

37

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATTCAATAGT TTACAGGCAA GTGCTACTGA GTACA

35

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TTGGCTACGC TTGGGCTATG GTAGTAGTTA TAGTT

35

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGTGCTACCA TAGGGATTAA ATTATTCAAA AAGTT

35

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TACGAGCAAG GCTTCTTA

18

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AGCTTAAGAA GCCTTGCTCG TAAACTTTTT GAATAATTT

39

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AATCCCTATG GTAGCACCAA CTATAACTAC TACCAT

36

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGCCCAAGCG TAGCCAATGT ACTCAGTAGC ACTTG

35

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CCTGTAAACT ATTGAATGCA GCCTTAGCAG GGTC

34

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ATCGCCTTCA GCCTAG

16

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CTCGAATTCG TACATCCTGG TCATAGC

27

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CATTTTTGCA GATGGCTTAG A

21

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TAGCATTAAAC GTCCAATA

18

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATATATTTTA GTAAGCTTCA TCTTCT

26

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GACAAAGAAC GCGTGAAAAC TTT

23

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GCGGGCCTCT TCGCTATTGC TTAAGAAGCC TTGCT

35

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 48 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TTCAGCCTAG GATCCGCCGA GCTCTCCTAC CTGCGAATTC GTACATCC

48

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TGGATTATAC TTCTAAATAA TGGA

24

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TAACACTCAT TCCGGATGGA ATTCTGGAGT CTGGGT

36

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AATTCGCCAA GGAGACAGTC AT

22

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

AATGAAATAC CTATTGCCTA CGGCAGCCGC TGGATTGTT

39

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATTACTCGCT GCCCAACCAG CCATGGCCGA GCTCGTGAT

39

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GACCCAGACT CCAGATATCC AACAGGAATG AGTGTTAAT

39

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TCTAGAACGC GTC

13

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

ACGTGACGCG TTCTAGAATT AACACTCATT CCTGT

35

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TGGATATCTG GAGTCTGGGT CATCACGAGC TCGGCCATG

39

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GCTGGTTGGG CAGCGAGTAA TAACAATCCA GCGGCTGCC

39

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GTAGGCAATA GGTATTTTCAT TATGACTGTC CTTGGCG

37

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TGACTGTCTC CTTGGCGTGT GAAATTGTTA

30

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

TAACACTCAT TCCGGATGGA ATTCTGGAGT CTGGGT

36

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CAATTTTATC CTAAATCTTA CCAAC

25

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CATTTTGTGCA GATGGCTTAG A

21

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CGAAAGGGGG GTGTGCTGCA A

21

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

TAGCATTAAAC GTCCAATA

18

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 43 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

AAACGACGGC CAGTGCCAAG TGACGCGTGT GAAATTGTTA TCC

43

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 43 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GGCGAAAGGG AATTCTGCAA GGCGATTAAG CTTGGGTAAC GCC

43

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GGCGTTACCC AAGCTTTGTA CATGGAGAAA ATAAAG

36

## (2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 42 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

TGAAACAAAG CACTATTGCA CTGGCACTCT TACCGTTACC GT

42

## (2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 42 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

TACTGTTTAC CCCTGTGACA AAAGCCGCCC AGGTCCAGCT GC

42

## (2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 44 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TCGAGTCAGG CCTATTGTGC CCAGGGATTG TACTAGTGGA TCCG

44

## (2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 38 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

TGGCGAAAGG GAATTCGGAT CCACTAGTAC AATCCCTG

38

## (2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

GGCACAATAG GCCTGACTCG AGCAGCTGGA CCAGGGCGGC TT

42

## (2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

TTGTCACAGG GGTAACAGT AACGGTAACG GTAAGTGTGC CA

42

## (2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GTGCAATAGT GCTTTGTTTC ACTTTATTTT CTCCATGTAC AA

42

## (2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

TAACGGTAAG AGTGCCAGTG C

21

## (52) INFORMATION FOR SEQ ID NO:52:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 68 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ix) FEATURE:
  - (A) NAME/KEY: misc\_difference
  - (B) LOCATION: replace(25, "")
  - (D) OTHER INFORMATION: /note= "M REPRESENTS AN EQUAL MIXTURE OF A AND C AT THIS LOCATION AND AT LOCATIONS 28, 31, 34, 37, 40, 43, 46 & 49"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

AGCTCCCGGA TGCCTCAGAA GATGMNNMNN MNNMNNMNNM NNMNNMNNMN NGGCTTTTGC 60  
CACAGGGG 68

- (2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 54 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ix) FEATURE:
  - (A) NAME/KEY: misc\_difference
  - (B) LOCATION: replace(17, "")
  - (D) OTHER INFORMATION: /note= "M REPRESENTS AN EQUAL MIXTURE OF A AND C AT THIS LOCATION AND AT LOCATIONS 20, 23, 26, 29, 32, 35, 38, 41, 44 & 50"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CAGCCTCGGA TCCGCCMNNM NNMNNMNNMN NMNNMNNMNN MNNMNNATGM GAAT 54

- (2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GGTAAACAGT AACGGTAAGA GTGCCAG 27

- (2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

GGGCTTTTGC CACAGGGGT

19

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 63 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

AGGGTCATCG CCTTCAGCTC CGGATCCCTC AGAAGTCATA AACCCCCCAT AGGCTTTTGC

60

CAC

63

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 47 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

TCGCCTTCAG CTCCCGGATG CCTCAGAAGC ATGAACCCCC CATAGGC

47

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 25 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

CAATTTTATC CTAAATCTTA CCAAC

25

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 21 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single



(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

GCCTTCAGCC TCGGATCCGC C

21

(2) INFORMATION FOR SEQ ID NO:60:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

CGGATGCCTC AGAAGCCCCN N

21

(2) INFORMATION FOR SEQ ID NO:61:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

CGGATGCCTC AGAAGGGCTT TTGCCACAGG

30